Benchmarking proximal methods acceleration enhancements for CS-acquired MR image analysis reconstruction

Zaccharie Ramzi, Philippe Ciuciu, Jean-Luc Starck

Benchmarking proximal methods acceleration enhancements for CS-acquired MR image analysis reconstruction

Zaccharie Ramzi
CEA - Neurospin & Cosmostat
INRIA - Parietal
Gif-sur-Yvette, France
Email: zaccharie.ramzi@cea.fr

Philippe Ciuciu
CEA - Neurospin
INRIA - Parietal
Gif-sur-Yvette, France
Email: philippe.ciuciu@cea.fr

Jean-Luc Starck
CEA - Cosmostat
Gif-sur-Yvette, France
Email: jstarck@cea.fr

I. Problem statement

Magnetic resonance imaging (MRI) is one of the most powerful imaging techniques for examining the human body since it allows early and accurate diagnosis of pathologies [1]. Although high magnetic field systems ($\geq$ 3 Tesla) enable increased spatial resolution, long scan times (i.e. 15 min for high-resolution (HR) imaging around 500 um isotropic [2]) and motion sensitivity continue to impede the exploitation of HR-MRI. To circumvent that problem, Compressed Sensing (CS) was introduced, among other techniques, to reduce the acquisition time, taking advantage of the structure of MR images [3]. However, the time gained on acquisition has been lost on reconstruction as sparse recovery amounts to iteratively solving a linear inverse problem, formalized as follows:

$$\underset{\boldsymbol{x} \in \mathbb{C}^{n \times n}}{\arg\min}\, \mathcal{J}(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{F}_\Omega \boldsymbol{x}\|_2^2 + \lambda \|\boldsymbol{\Psi} \boldsymbol{x}\|_1 \tag{1}$$

where $\boldsymbol{x}$ denotes the sought complex-valued MR image, n is its dimension, F_Ω is the Fourier operator, possibly non-uniform and undersampled over the non-Cartesian set Ω, $\boldsymbol{y}$ the MR Fourier data also called k-space samples. Scalar parameter λ refers to the regularization parameter and $\boldsymbol{\Psi}$ to the wavelet decomposition operator as the MR image is assumed to be sparse (at least compressible) in the wavelet basis [3].

The way to solve problem (1) is typically with a proximal gradient method such as FISTA [4] or POGM [5], or a primal/dual method (ADMM [6] or Condat-Vu [7] algorithms). These methods require a *forward* pass and a *backward* pass (application of the adjoint) of the two operators (Fourier and wavelets) at each iteration and are therefore very time consuming, especially in large dimensions (2 min for a 512×512 - 500 µm in plane resolution - slice of a given brain phantom on a machine with 8 cores). The motivation to reduce the time required for image reconstruction is the following: the MRI physician needs to analyze the reconstructed image to know if there has been some movement causing some motion artifacts [8], and therefore take the decision to rerun the exam [9]. Thus, decreasing reconstruction time would mean reducing the overall duration of the MRI exam. This would also help approaching the goal of real-time MRI, which is useful for monitoring cardiovascular procedures for example [10]. This becomes substantially important when going from 2D to 3D to 4D (3D + time or contrast) with very high-resolution, namely 1204x1204 or 250 µm in plane resolution at ultra-high fields ($\geq$ 7 Tesla).

The goal of this work was to focus on the different techniques listed in [11] and benchmark their speed and memory load against the vanilla FISTA, the Condat-Vu algorithm and the POGM' algorithm (i.e. POGM combined with adaptive restart) on the problem (1). The most promising algorithm, greedy FISTA, applies the following modifications to the vanilla FISTA (using notation from [4]):

- It forces the $\frac{t_{k-1}-1}{t_k}$ term to stick to 1, providing a constant momentum.
- It restarts the algorithm when a gradient-like criterion is met to prevent rippling.
- It allows having a larger constant step size by reducing it when a safeguard criterion is met.

These algorithms are all implemented in the `Python` library `modopt` [12] for the sake of reproducibility.

The criterion used to assert convergence is the objective function (although the results do extend to problem-dependent criteria like normalized residual mean square error, etc.). Our analysis was performed on the 2D MR phantom image shown in Fig. 2a, which was non-uniformly Fourier transformed to pick up the k-space samples depicted in Fig. 2b when using the random variable density (VD) sampling scheme. We also tried the same analysis with the SPARKLING sampling scheme [13], which is a physically feasible accelerated sampling scheme. The data obtained with this scheme is shown in Fig. 2c. The orthogonal Daubechies 4 was chosen for the wavelet basis. This made the proximal of the regularisation term computable in closed form [14].

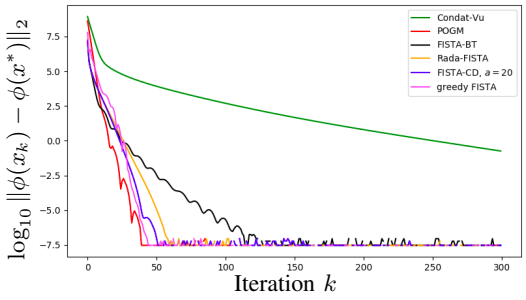
II. Results

The results presented in Fig. 1 show that the greedy FISTA compares to POGM' in terms of time (only a few iterations more needed for greedy FISTA to converge). However, it theoretically uses twice less memory. This might be of a huge importance when scaling these algorithms to 3D parallel imaging. In the latter context, 32 to 64 k-space are collected simultaneously over multiple receivers, each of them going up to 512^3 in dimension. The results also confirm the interest of using greedy FISTA compared to vanilla FISTA: it is 3 times faster than its ancestors.

III. Related works

In [15], the benchmark was done on the Forward Backward algorithm, the vanilla FISTA and POGM. The problem studied was slightly more ambitious as it involved parallel imaging, but this comparison did not embody the most recent accelerations of FISTA. In [16], an overview of multiple concurrent formulations of the CS-MRI problem was presented. These formulations depart from Eq. (1) in the way sparsity is promoted (e.g. analysis vs synthesis regularization, total variation, etc.) and the corresponding algorithms to solve them were summarized. Finally, the original paper presenting the FISTA enhancements [11] also benchmarks them but on other applications and not against a totally different algorithm (namely here Condat-Vu and POGM').

(a) Non-uniform $\boldsymbol{F}_\Omega$ using a non-realistic independent VD sampling scheme



(b) Non-uniform $\boldsymbol{F}_\Omega$, physically feasible scheme with an acceleration factor (AF) of 15 (undersampling of 40%)

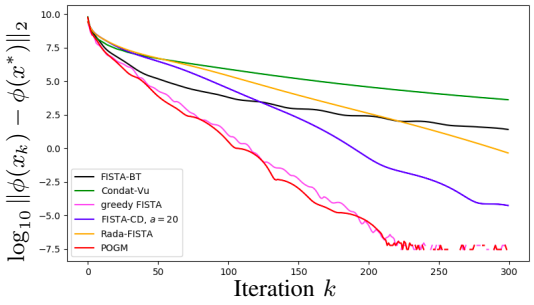
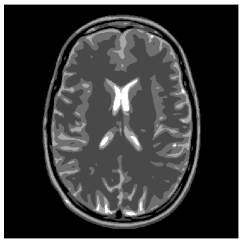
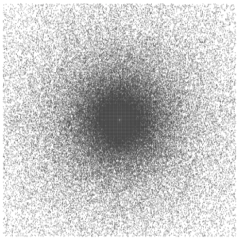


Fig. 1: Comparison of the convergence speed of different algorithms to minimize $\mathcal{J}(\boldsymbol{x})$. Parameters used for: 1) FISTA-CD (by Chambolle and Dossal [17]): $a = 20$; 2) Rada-FISTA: $p = \frac{1}{30}$, $q = \frac{1}{10}$, $\xi = 0.96$; 3) greedy FISTA: $\xi = 0.96$, $S = 1.1$; 4) POGM': $\overline{\sigma} = 0.96$. 5) Condat-Vu: $\rho = 1$, $\sigma = 10$. Different support Ω were tested in the non-uniform Fourier operator $\boldsymbol{F}_\Omega$. The reference is the original implementation proposed by Beck and Teboulle (FISTA-BT). Greedy-FISTA and POGM' converge faster.

(a) 2D MRI phantom (512×512)



(b) VD sampling (25% undersampling)



(c) SPARKLING scheme (AF=15)

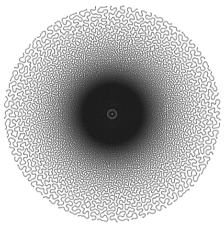


Fig. 2: True MR image and non-uniform k-space data used to benchmark algorithms.

References

[1] A. G. van der Kolk et al. "Clinical applications of 7T MRI in the brain". In: *European journal of radiology* 82.5 (2013), pp. 708–718.

[2] C. Lazarus et al. "3D SPARKLING trajectories for high-resolution T2*-weighted Magnetic Resonance imaging". In: (2019).

[3] M. Lustig, D. Donoho, and J. M. Pauly. "Sparse MRI: The application of compressed sensing for rapid MR imaging". In: *Magnetic Resonance in Medicine* 58.6 (2007), pp. 1182–1195.

[4] A. Beck and M. Teboulle. "A fast iterative shrinkage-thresholding algorithm for linear inverse problems". In: *SIAM Journal on Imaging Sciences* 2.1 (2009), pp. 183–202.

[5] D. Kim and J. A. Fessler. "Adaptive restart of the optimized gradient method for convex optimization". In: *Journal of Optimization Theory and Applications* 178.1 (2018), pp. 240–263.

[6] S. Boyd et al. "Distributed optimization and statistical learning via the alternating direction method of multipliers". In: *Found. and Trends in Mach. learn.* 3.1 (2011), pp. 1–122.

[7] L. Condat. "A Primal–Dual Splitting Method for Convex Optimization Involving Lipschitzian, Proximable and Linear Composite Terms". In: *Journal of Optimization Theory and Applications* 158.2 (2013), pp. 460–479.

[8] M. Zaitsev, J. Maclaren, and M. Herbst. "Motion artifacts in MRI: a complex problem with many partial solutions". In: *Journal of Magnetic Resonance Imaging* 42.4 (2015), pp. 887–901.

[9] A. S. Health. *Clinical Appropriateness Guidelines: Advanced Imaging*. 2017. URL: https://www.aimspecialtyhealth.com/PDF/Guidelines/2017/Sept05/AIM_Guidelines.pdf (visited on 03/26/2019).

[10] K. A. Horvath et al. "Real-time magnetic resonance imaging guidance for cardiovascular procedures". In: *Seminars in thoracic and cardiovascular surgery*. Vol. 19. 4. Elsevier. 2007, pp. 330–335.

[11] J. Liang and C.-B. Schönlieb. "Faster FISTA". In: *arXiv e-prints*, arXiv:1807.04005 (2018), arXiv:1807.04005.

[12] S. Farrens. *ModOpt: Modular Optimisation tools for solving inverse problems in Python*. [Online; accessed ¡today¿]. 2017–.

[13] C. Lazarus et al. "SPARKLING: Novel non-Cartesian sampling schemes for accelerated 2D anatomical imaging at 7T using compressed sensing". In: *25th annual meeting of the International Society for Magnetic Resonance Imaging*. Honolulu, HW, USA, 2017.

[14] P. L. Combettes and V. R. Wajs. "Signal recovery by proximal forward-backward splitting". In: *Multiscale Modeling & Simulation* 4.4 (2005), pp. 1168–1200.

[15] L. El Gueddari et al. "Self-calibrating nonlinear reconstruction algorithms for variable density sampling and parallel reception MRI". In: *10th IEEE Sensory Array and Multichannel (SAM) signal processing workshop*. Sheffield, UK, 2018, pp. –.

[16] J. A. Fessler. "Optimization methods for MR image reconstruction". In: *arXiv preprint arXiv:1903.03510* (2019).

[17] A. Chambolle and C. Dossal. "On the convergence of the iterates of" FISTA"". In: *Journal of Optimization Theory and Applications* 166.3 (2015), p. 25.